

Individual Proposal

Method 1: Confidence Interval Analysis

Research Focus and Question

The focus of my research is to estimate the overall level of life satisfaction among respondents. My research question is: “What is the average level of life satisfaction among the surveyed population?”

Selected Variables and Data Collection

I will use the variable Life Satisfaction Score from the CSCS dataset. This numeric variable measures an individual’s subjective well-being and is crucial for understanding the general happiness level within the sample.

Data Visualization

To visualize the distribution of life satisfaction scores, I will use a Histogram. This visualization will help show the concentration and spread of the data, and identify any significant outliers, making it easier to judge whether the data is suitable for confidence interval analysis.

Assumptions for Analysis

For the confidence interval to be valid, we assume that the data is randomly sampled and independent. These assumptions ensure that the confidence interval accurately reflects the population.

Method of Analysis

I will calculate a Confidence Interval for the mean level of life satisfaction. The process includes calculating the sample mean and standard error, then using the formula “sample mean \pm critical value * standard error” to find the upper and lower bounds of the interval. This method will allow me to estimate a range for the average life satisfaction in the population.

Potential Outcomes

A narrow confidence interval at a set confidence level (e.g., 95%) would suggest an accurate estimate of life satisfaction. Conversely, a wide interval would indicate greater variability in life satisfaction levels. A high mean with a narrow interval suggests that the population is generally satisfied, whereas a low mean and wider interval suggest lower, more varied satisfaction.

Relevance to Course Objectives

This method is like the confidence interval analysis we practiced in Week 4. In that tutorial, we used confidence intervals to estimate the mean of a sample, helping us understand how sample data can infer population levels. Here, I apply the method to life satisfaction, providing insight into the overall happiness of the population, which aligns with the project's goal of exploring relationships between social connections and personal well-being.

Method 2: Hypothesis Testing

Research Focus and Question

My second focus is to determine if social support has a significant effect on an individual's subjective well-being. The research question is: "Does social support significantly influence life satisfaction?"

Selected Variables and Data Collection

I will use two variables from the CSCS dataset: Life Satisfaction Score and Social Support. Life satisfaction is a numeric variable measuring subjective well-being, and social support is a binary variable coded as 0 for no support and 1 for sufficient support. These variables allow us to test if support significantly impacts life satisfaction.

Data Visualization

To visualize differences in life satisfaction across social support groups, I will use a Box Plot. This visualization displays the median, quartiles, and outliers for life satisfaction in each group, highlighting any differences between the support and non-support groups.

Assumptions for Analysis

We assume that the samples are independent, and that life satisfaction scores in each group approximately follow a normal distribution. These assumptions are necessary for valid hypothesis testing.

Method of Analysis

I will use a Two-Sample t-Test to evaluate the effect of social support on life satisfaction. I will set up a null hypothesis: social support does not significantly affect life satisfaction, and an alternative hypothesis: social support has a significant effect. If the p-value from this test is below the significance level (e.g., 0.05), I will reject the null hypothesis, suggesting a significant impact of social support on happiness.

Potential Outcomes

If the test reveals a significant difference, with a p-value below the threshold, this suggests that social support positively impacts life satisfaction. If the p-value is high, we fail to reject the null hypothesis, indicating that social support may not significantly influence happiness in this sample. These results could reveal the role of social support in improving subjective well-being.

Relevance to Course Objectives

This hypothesis testing approach is like the analysis conducted in Week 5, where we examined group differences using hypothesis testing. That tutorial helped us understand how statistical significance can determine the difference between groups. Here, I apply the same method to test the relationship between social support and life satisfaction, which is essential for the project's goal of exploring connections between social relationships and well-being.

Method 3: Regression Analysis

Research Focus and Question

My final focus is to assess how social interaction frequency affects subjective well-being. The research question is: "Does the frequency of social interactions significantly influence life satisfaction?"

Selected Variables and Data Collection

I will use Life Satisfaction Score as the dependent variable and Social Interaction Frequency as the independent variable. Life satisfaction measures subjective well-being, while social interaction frequency captures the extent of social engagement. These variables will help identify the potential impact of social interactions on happiness.

Data Visualization

To visualize the relationship between social interaction frequency and life satisfaction, I will use a Scatter Plot. The scatter plot will show any trends and allow us to observe whether life satisfaction increases with more frequent interactions.

Assumptions for Analysis

For the regression model to be valid, we assume a linear relationship between the variables, independent samples, homoscedasticity of residuals, and normally distributed errors. These assumptions ensure the accuracy of the regression analysis.

Method of Analysis

I will use Simple Linear Regression to quantify the impact of social interaction frequency on life satisfaction. I will set up a model as follows: $\text{Life Satisfaction} = \text{Intercept} + \text{Slope} * \text{Social Interaction Frequency} + \text{Residual}$. The least squares method will estimate the regression coefficients, and the model will indicate how life satisfaction is affected by social interaction frequency.

Potential Outcomes

I expect the regression to reveal a positive relationship, where increased interaction frequency correlates with higher life satisfaction. A significant model suggests that social interaction frequency is a strong predictor of happiness, whereas a non-significant model may indicate a minor impact. These findings will provide insights into how social interactions contribute to well-being.

Relevance to Course Objectives

This regression analysis aligns with the concepts introduced in Week 6, where we explored linear regression to understand the relationship between variables. In that tutorial, we learned to use simple linear regression models to predict outcomes based on one independent variable. Here, I apply this knowledge to assess the impact of social interaction frequency on life satisfaction. By examining how social engagement influences well-being, this analysis directly supports the project's objective of investigating the connections between social relationships and personal happiness.